GenCore version 5.1.4_p5_4578


OM protein - protein search, using sw model

Run on: March 24, 2003, 15:45:24 ; Search time 55.1515 Seconds
(without alignments)
628.181 Million cell updates/sec

Title: US-09-988-971-2_COPY_2_261
Perfect score: 1346
Sequence: 1 GSLPSRRKSLPSPSLSSSVQ.....RESLSFYISLNDEAVSLDDA 260

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1346 | 100.0 | 261 | 23 | AAO15457 | Human modulator of |
| 2 | 1342 | 99.7 | 261 | 23 | AAU91308 | Human protein NOV1 |
| 3 | 1273 | 94.6 | 248 | 21 | AAB42993 | Human ORFX ORF2757 |
| 4 | 1027 | 76.3 | 259 | 23 | AAO15456 | Mouse modulator of |
| 5 | 933.5 | 69.4 | 210 | 23 | AAO15458 | Mouse modulator of |
| 6 | 474.5 | 35.3 | 315 | 22 | AAU31072 | Novel human secret |
| 7 | 364.5 | 27.1 | 505 | 22 | AAB99332 | Human tyrosine kin |
| 8 | 352 | 26.2 | 509 | 21 | AAY49420 | PKA substrate, Src |
| 9 | 344 | 25.6 | 508 | 21 | AAB37700 | Human lymphocyte k |
| 10 | 342 | 25.4 | 70 | 22 | ABG05994 | Novel human diagno |
| 11 | 336 | 25.0 | 517 | 22 | ABB57957 | Drosophila melanog |
| 12 | 321 | 23.8 | 541 | 23 | AAU74614 | Perinuclear theca |
| 13 | 319.5 | 23.7 | 543 | 22 | ABG10302 | Novel human diagno |
| 14 | 319.5 | 23.7 | 543 | 22 | AAB84663 | Amino acid sequenc |
| 15 | 317.5 | 23.6 | 543 | 20 | AAY24421 | Human yes1 protein |
| 16 | 315 | 23.4 | 496 | 20 | AAY29668 | Human src-family k |
| 17 | 314 | 23.3 | 496 | 22 | AAU08730 | Xenopus laevis src |
| 18 | 314 | 23.3 | 496 | 22 | AAU08734 | Xenopus laevis src |
| 19 | 314 | 23.3 | 496 | 22 | AAU08735 | Xenopus laevis src |
| 20 | 310.5 | 23.1 | 551 | 22 | ABG22264 | Novel human diagno |
| 21 | 290.5 | 21.6 | 533 | 21 | AAY44447 | Wild-type chicken |
| 22 | 290.5 | 21.6 | 533 | 21 | AAY44449 | Mutant chicken c-S |
| 23 | 290.5 | 21.6 | 533 | 22 | AAB84661 | Amino acid sequenc |
| 24 | 290.5 | 21.6 | 552 | 22 | ABB57777 | Drosophila melanog |
| 25 | 290.5 | 21.6 | 552 | 22 | ABB66713 | Drosophila melanog |
| 26 | 288.5 | 21.4 | 533 | 14 | AAR39705 | Chicken pp60 c-src |
| 27 | 286 | 21.2 | 502 | 23 | AAE21689 | Fugu rubripes lymp |
| 28 | 286 | 21.2 | 533 | 21 | AAY44451 | Mutant chicken c-S |
| 29 | 285.5 | 21.2 | 251 | 21 | AAY44450 | Mutant chicken c-S |
| 30 | 280.5 | 20.8 | 536 | 14 | AAR39706 | Human pp60 c-src p |
| 31 | 280.5 | 20.8 | 536 | 23 | AAU78678 | Human SH2/SH3 doma |
| 32 | 277.5 | 20.6 | 542 | 23 | ABB97339 | Novel human protei |
| 33 | 266 | 19.8 | 134 | 17 | AAW03982 | DET1-DET2-spacer-e |
| 34 | 266 | 19.8 | 134 | 17 | AAW02120 | DET1-DET2-spacer-e |
| 35 | 266 | 19.8 | 134 | 18 | AAW11286 | DET1-DET2-spacer-e |
| 36 | 266 | 19.8 | 134 | 18 | AAW19624 | Human lck SH2 doma |
| 37 | 264 | 19.6 | 101 | 18 | AAW31184 | Human p56-lck prot |
| 38 | 262 | 19.5 | 224 | 18 | AAW14788 | FKBP-LCK:SH2 fusio |
| 39 | 262 | 19.5 | 224 | 20 | AAW96823 | A fusion protein o |
| 40 | 258.5 | 19.2 | 102 | 16 | AAR72090 | Lck SH2 region. N |
| 41 | 256 | 19.0 | 565 | 22 | ABG23778 | Novel human diagno |
| 42 | 242 | 18.0 | 417 | 12 | AAR14201 | (Beta-galactosidas |
| 43 | 238 | 17.7 | 94 | 20 | AAY29670 | Human src-family k |
| 44 | 238 | 17.7 | 94 | 22 | AAU08732 | Src-family kinase |
| 45 | 237 | 17.6 | 117 | 17 | AAW03986 | SH2 domain from hu |

ALIGNMENTS

RESULT 1
AAO15457
ID AAO15457 standard; Protein; 261 AA.
XX
AC AAO15457;
XX
DT 03-OCT-2002 (first entry)
XX
DE Human modulator of antigen receptor signalling (MARS) protein.
XX
KW Human; gene therapy; modulator of antigen receptor signalling; MARS;
KW tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS Homo sapiens.
XX
PN WO200242452-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001WO-CA01662.
XX
PR 27-NOV-2000; 2000CA-2324663.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI Mcglade JC, Loreto MP;
XX
DR WPI; 2002-566564/60.
DR N-PSDB; AAL44089.
XX
PT New isolated modulator of antigen receptor signaling protein or its

PT fragment, useful for treating malignant disorders such as myeloid
PT malignancies, autoimmune disorders and myeloproliferative disorders -
XX
PS Claim 7; Fig 9A; 110pp; English.
XX
CC The invention comprises the amino acid and coding sequences of modulator
CC of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC putative tumour suppressor gene and exhibits structural and sequence
CC similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC protein sequences of the invention are useful for the treatment of
CC myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC disorders, immunosuppression, myeloproliferative disorders and
CC malignancies related to the de-regulation of tyrosine kinases (e.g.
CC breast cancer). The present amino acid sequence represents a human MARS
CC protein.
XX
SQ Sequence 261 AA;

Query Match 100.0%; Score 1346; DB 23; Length 261;
Best Local Similarity 100.0%; Pred. No. 5.9e-130;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 61

Qy     61 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121

Qy    121 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELAD 181

Qy    181 DICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 DICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGL 241

Qy    241 RESLSFYISLNDEAVSLDDA 260
          ||||||||||||||||||||
Db    242 RESLSFYISLNDEAVSLDDA 261
```

RESULT 2
AAU91308
ID AAU91308 standard; Protein; 261 AA.
XX
AC AAU91308;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human protein NOV13.
XX
KW Human; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;
KW uterus cancer; immune response; graft-versus-host disease;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy.
XX
OS Homo sapiens.
XX
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US26510.
XX
PR 25-AUG-2000; 2000US-228191P.
PR 08-FEB-2001; 2001US-267300P.
PR 20-FEB-2001; 2001US-269961P.
PR 20-MAR-2001; 2001US-277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
DR N-PSDB; ABK61465.
XX
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers -
XX
PS Claim 3; Page 98; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature
CC form of NOVX, a NOVX variant (differing by no more than 15%), the
CC nucleotide encoding NOVX (or its complement, fragment or variant).
CC NOVX is NOV1-14, 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic
CC acid encoding it and antibody against it, are useful for treating or
CC preventing (e.g. by gene therapy) a NOVX-associated disorder in humans,
CC e.g. cardiomyopathy, atherosclerosis, a disorder related to cell signal
CC processing and metabolic pathway modulation, diabetes or cancers. The
CC NOVX polypeptide and nucleic acids are also useful for determining the
CC presence of predisposition to the diseases. The NOVX nucleic acid and
CC polypeptide are especially useful in therapeutic or prophylactic
CC applications for disorders associated with aberrant NOVX expression or
CC activity, e.g. cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or
CC uterus cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy
CC for treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of
CC various human disorders, as well as in diagnostic applications. The
CC present sequence represents a NOVX protein.
XX
SQ Sequence 261 AA;

Query Match 99.7%; Score 1342; DB 23; Length 261;
Best Local Similarity 99.6%; Pred. No. 1.5e-129;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 61

Qy     61 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 120
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     62 VSEDGDWWTVLSEVSGREYNIPSVHVGKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121

Qy    121 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELAD 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELAD 181

Qy    181 DICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 DICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGL 241

Qy    241 RESLSFYISLNDEAVSLDDA 260
          ||||||||||||||||||||
Db    242 RESLSFYISLNDEAVSLDDA 261
```

RESULT 3
AAB42993
ID AAB42993 standard; Protein; 248 AA.
XX
AC AAB42993;
XX
DT 08-FEB-2001 (first entry)

```
XX
DE   Human ORFX ORF2757 polypeptide sequence SEQ ID NO:5514.
XX
KW   Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW   vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW   anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW   immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW   hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW   antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW   antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW   neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW   cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW   cholesterol ester storage; systemic lupus erythematosus; infection;
KW   severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW   allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW   bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW   thrombosis; contraceptive.
XX
OS   Homo sapiens.
XX
PN   WO200058473-A2.
XX
PD   05-OCT-2000.
XX
PF   31-MAR-2000; 2000WO-US08621.
XX
PR   31-MAR-1999;   99US-0127607.
PR   02-APR-1999;   99US-0127636.
PR   05-APR-1999;   99US-0127728.
PR   30-MAR-2000; 2000US-0540763.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2000-602362/57.
DR   N-PSDB; AAC77202.
XX
PT   Novel nucleic acids and peptides derived from open reading frame X,
PT   useful for treating e.g. cancers, proliferative disorders,
PT   neurodegenerative disorders and cardiovascular disease -
XX
PS   Claim 11; Page 4693-4694; 5507pp; English.
XX
CC   AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC   which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC   sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC   antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
CC   osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
CC   immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC   antidiabetic; hypotensive; dermatological; immunosuppressive;
CC   antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
CC   antithyroid; and antianaemic. The sequences can be used for determining
CC   the presence of or predisposition to, or preventing or treating
CC   pathological conditions associated with an ORFX-associated disorder. The
CC   nucleic acids can be used to express ORFX proteins in gene therapy
CC   vectors. The proteins and nucleic acids may be used to treat cancers,
CC   proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC   graft vs host disease, cardiovascular disease, diabetes mellitus,
CC   hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC   erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC   bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC   allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC   nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC   coagulation; to inhibit thrombosis; and as a contraceptive.
XX
SQ   Sequence   248 AA;

  Query Match            94.6%;  Score 1273;  DB 21;  Length 248;
  Best Local Similarity  99.2%;  Pred. No. 1.8e-122;
  Matches  245;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     14 SLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTIVSEDGDWWTVLSE 73
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db      2 SLSSSVQGQGPVTMEAERSKATAAALGSFPAGGPAELSLRLGEPLTIVSEDGDWWTVLSE 61

Qy     74 VSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSV 133
           |||||||||||||||||||||||||||:||||||||||||||||||||||||||||||||
Db     62 VSGREYNIPSVHVAKVSHGWLYEGLSREKAEDLLLLPGNPGGAFLIRESQTRRGSYSLSV 121

Qy    134 RLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPCVLQ 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 RLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPCVLQ 181

Qy    194 RAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGLRESLSFYISLNDE 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 RAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGEESLLSEGLRESLSFYISLNDE 241

Qy    254 AVSLDDA 260
           |||||||
Db    242 AVSLDDA 248


RESULT 4
AAO15456
ID   AAO15456 standard; Protein; 259 AA.
XX
AC   AAO15456;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Mouse modulator of antigen receptor signalling (MARS) protein.
XX
KW   Mouse; gene therapy; modulator of antigen receptor signalling; MARS;
KW   tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW   myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW   immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS   Mus sp.
XX
PN   WO200242452-A2.
XX
PD   30-MAY-2002.
XX
PF   26-NOV-2001; 2001WO-CA01662.
XX
PR   27-NOV-2000; 2000CA-2324663.
XX
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Mcglade JC,  Loreto MP;
XX
DR   WPI; 2002-566564/60.
DR   N-PSDB; AAL44087.
XX
PT   New isolated modulator of antigen receptor signaling protein or its
PT   fragment, useful for treating malignant disorders such as myeloid
PT   malignancies, autoimmune disorders and myeloproliferative disorders  -
XX
PS   Claim 7; Fig 1A; 110pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of modulator
CC   of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC   putative tumour suppressor gene and exhibits structural and sequence
CC   similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC   protein sequences of the invention are useful for the treatment of
CC   myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC   disorders, immunosuppression, myeloproliferative disorders and
CC   malignancies related to the de-regulation of tyrosine kinases (e.g.
CC   breast cancer). The present amino acid sequence represents a mouse MARS
CC   protein.
XX
SQ   Sequence   259 AA;

  Query Match            76.3%;  Score 1027;  DB 23;  Length 259;
  Best Local Similarity  79.7%;  Pred. No. 4.1e-97;
```

```
  Matches  208;  Conservative   16;  Mismatches   33;  Indels    4;  Gaps    3;

Qy      1 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 60
          ||| || |:  ||| |||   | ||:|: || | ||||||||||  | ||||||||||
Db      2 GSLSSRGKT-SSPSPSSSGPDQEPVSMQPERHKVTAVALGSFPAGEQARLSLRLGEPLTI 60

Qy     61 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 120
          :|||||||| |||||||::|||:||||:|||||||||||||||||||||||||||||||
Db     61 ISEDGDWWTVQSEVSGREYHMPSVYVAKVAHGWLYEGLSREKAEELLLLPGNPGGAFLIR 120

Qy    121 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELAD 180
          |||||| |||||||||||||||||||  |||||||||||||||| |||:|||||||
Db    121 ESQTRRGCYSLSVRLSRPASWDRIRHYRIQRLDNGWLYISPRLTFPSLHALVEHYSELAD 180

Qy    181 DICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEA-ATGEESLLSEG 239
          |||  |:||||||: ||||||| | ||||  : ||||:|| |||| || |:|| ||||||
Db    181 GICCPLREPCVLQKLGPLPGKDTPPPVTVPTSSLNWKKLDRSLLFLEAPASGEASLLSEG 240

Qy    240 LRESLSFYISLNDEAVSLDDA 260
          |||||| |||| ::   ||||
Db    241 LRESLSSYISLAED--PLDDA 259


RESULT 5
AAO15458
ID   AAO15458 standard; Protein; 210 AA.
XX
AC   AAO15458;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Mouse modulator of antigen receptor signalling short isoform protein.
XX
KW   Mouse; gene therapy; modulator of antigen receptor signalling; MARS;
KW   tumour suppressor gene; Scr-like adaptor protein; SLAP;
KW   myeloid malignancy; acute myelogenous leukaemia; autoimmune disorder;
KW   immunosuppression; myeloproliferative disorder; breast cancer.
XX
OS   Mus sp.
XX
PN   WO200242452-A2.
XX
PD   30-MAY-2002.
XX
PF   26-NOV-2001; 2001WO-CA01662.
XX
PR   27-NOV-2000; 2000CA-2324663.
XX
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Mcglade JC,  Loreto MP;
XX
DR   WPI; 2002-566564/60.
DR   N-PSDB; AAL44090.
XX
PT   New isolated modulator of antigen receptor signaling protein or its
PT   fragment, useful for treating malignant disorders such as myeloid
PT   malignancies, autoimmune disorders and myeloproliferative disorders  -
XX
PS   Claim 8; Page 78; 110pp; English.
XX
CC   The invention comprises the amino acid and coding sequences of modulator
CC   of antigen receptor signalling (MARS) proteins. The MARS protein is a
CC   putative tumour suppressor gene and exhibits structural and sequence
CC   similarity to the Scr-like adaptor protein (SLAP). The MARS DNA and
CC   protein sequences of the invention are useful for the treatment of
CC   myeloid malignancies (e.g. acute myelogenous leukaemia) autoimmune
CC   disorders, immunosuppression, myeloproliferative disorders and
CC   malignancies related to the de-regulation of tyrosine kinases (e.g.
CC   breast cancer). The present amino acid sequence represents a mouse MARS
CC   protein.
XX
SQ   Sequence  * 210 AA;

  Query Match             69.4%;  Score 933.5;  DB 23;  Length 210;
  Best Local Similarity   89.3%;  Pred. No. 1.3e-87;
  Matches  184;  Conservative    3;  Mismatches   12;  Indels    7;  Gaps    1;

Qy      1 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 61

Qy     61 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 VSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIR 121

Qy    121 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSE--- 177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 ESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSEGWP 181

Qy    178 ----LADDICCLLKEPCVLQRAGPLP 199
              |    ::    |:||| ||
Db    182 APWQGYTPTCDCAEDTTQLERAGQLP 207


RESULT 6
AAU31072
ID   AAU31072 standard; Protein; 315 AA.
XX
AC   AAU31072;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Novel human secreted protein #1563.
XX
KW   Human; vaccination; gene therapy; nutritional supplement;
KW   stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW   immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200179449-A2.
XX
PD   25-OCT-2001.
XX
PF   16-APR-2001; 2001WO-US08656.
XX
PR   18-APR-2000; 2000US-0552929.
PR   26-JAN-2001; 2001US-0770160.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-611725/70.
XX
PT   Nucleic acids encoding a range of human polypeptides, useful in genetic
PT   vaccination, testing and therapy -
XX
PS   Claim 20; Page 399; 765pp; English.
XX
CC   The invention relates to novel human secreted polypeptides. The
CC   polypeptides and antibodies to the polypeptides are useful for
CC   determining the presence of or predisposition to a disease associated
CC   with altered levels of polypeptide. The polypeptides are also useful for
CC   identifying agents (agonists and antagonists) that bind to them. Cells
CC   expressing the proteins are useful for identifying a therapeutic agent
CC   for use in treatment of a pathology related to aberrant expression or
CC   physiological interactions of the polypeptide. Vectors comprising
CC   the nucleic acids encoding the polypeptides and cells genetically
CC   engineered to express them are also useful for producing the proteins.
CC   The proteins are useful in genetic vaccination, testing and
CC   therapy, and can be used as nutritional supplements. They may be used to
CC   increase stem cell proliferation; to regulate haematopoiesis; and in
CC   bone, cartilage, tendon  and/or nerve tissue growth or regeneration;
CC   immune suppression and/or stimulation; as anti-inflammatory agents; and
```

```
CC   in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
CC   sequences of novel human secreted proteins of the invention.
XX
SQ   Sequence   315 AA;

  Query Match            35.3%;  Score 474.5;  DB 22;  Length 315;
  Best Local Similarity  39.7%;  Pred. No. 4.3e-40;
  Matches  104;  Conservative   46;  Mismatches   99;  Indels   13;  Gaps    4;

Qy       4 PSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTIVSE 63
Db      33 PGKKKEMGNSMKSTPAPAERPLPNPEGLDSDFLAVLSDYPSPDISPPIFRRGEKLRVISD 92

Qy      64 DGDWWTVLSEVSGREYNIPSVHVAKVSHG-WLYEGLSREKAEELLLLPGNPGGAFLIRES 122
Db      93 EGGWWKAISLSTGRESYIPGICVARVYHGLWLFEGLGRDKAEELLQLPDTKVGSFMIRES 152

Qy     123 QTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDI 182
Db     153 ETKKGFYSLSVR------HRQVKTYRIFRLPNNWYYISPRLTFQCLEDLVNHYSEVADGL 206

Qy     183 CCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWKELDSSLLFSEAATGE-----ESLLS 237
Db     207 CCVLTTPCLTQSTAAPAVRACSSPVTLRQKTVDWRRV-SRLQEDPEGTENPLGVVESLFS 265

Qy     238 EGLRESLSFYISLNDEAVSLDD 259
Db     266 YGLRESIASYLSLTSEDISSFD 287


RESULT 7
AAB99332
ID   AAB99332 standard; Protein; 505 AA.
XX
AC   AAB99332;
XX
DT   23-AUG-2001  (first entry)
XX
DE   Human tyrosine kinase Hck protein sequence SEQ ID NO:11.
XX
KW   Human; tyrosine kinase Hck binding protein; tyrosine kinase; Hck;
KW   tumour lethal factor; tumour necrosis factor alpha; apoptosis; HSB-1;
KW   Hck signal transduction; human immunodeficiency virus; HIV infection;
KW   anticancer.
XX
OS   Homo sapiens.
XX
PN   WO200132869-A1.
XX
PD   10-MAY-2001.
XX
PF   26-OCT-2000; 2000WO-JP07500.
XX
PR   29-OCT-1999;   99JP-0309957.
XX
PA   (SSSE ) SSP CO LTD.
XX
PI   Taniyama T,  Narita T;
XX
DR   WPI; 2001-316440/33.
XX
PT   New proteins which bind to human tyrosine kinase Hck for promotion of
PT   apoptosis and for the elucidation of the mechanism of Hck signal
PT   transduction  -
XX
PS   Example 1; Page 33-35; 45pp; Japanese.
XX
CC   The present invention describes a protein, designated HSB-1, which binds
CC   to human tyrosine kinase Hck. Also described are: (1) nucleic acids
CC   encoding the protein and its derivatives; (2) recombinant vectors
CC   containing the nucleic acids; and (3) host cells transformed by the
CC   vectors and expressing the protein. HSB-1 has cytostatic activity, binds
CC   tyrosine kinase, enhances tumour necrosis factor alpha and promotes
CC   apoptosis. HSB-1 proteins are used for the elucidation of the mechanism
CC   of Hck signal transduction and of the role of Hck in human
CC   immunodeficiency virus (HIV) infection. They can be used for the
CC   treatment of infections and other diseases with which Hck is associated.
CC   They promote the anticancer activity of tumour necrosis factor alpha.
CC   The present sequence represents the human tyrosine kinase Hck protein,
CC   which is used in an example from the present invention.
XX
SQ   Sequence   505 AA;

  Query Match            27.1%;  Score 364.5;  DB 22;  Length 505;
  Best Local Similarity  42.2%;  Pred. No. 1.9e-28;
  Matches   78;  Conservative   31;  Mismatches   69;  Indels    7;  Gaps    2;

Qy      11 PSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTIVSEDGDWWTV 70
Db      40 PGPNSHNS---NTPGIREAGSEDIIVVALYDYEAIHHEDLSFQKGDQMVVLEESGEWWKA 96

Qy      71 LSEVSGREYNIPSVHVAKV----SHGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRR 126
Db      97 RSLATRKEGYIPSNYVARVDSLETEEWFFKGISRKDAERQLLAPGNMLGSFMIRDSETTK 156

Qy     127 GSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLL 186
Db     157 GSYSLSVRDYDPRQGDTVKHYKIRTLDNGGFYISPRSTFSTLQELVDHYKKGNDGLCQKL 216

Qy     187 KEPCV 191
Db     217 SVPCM 221


RESULT 8
AAY49420
ID   AAY49420 standard; Protein; 509 AA.
XX
AC   AAY49420;
XX
DT   13-MAR-2000  (first entry)
XX
DE   PKA substrate, Src-family protein.
XX
KW   Protein kinase A; PKA; PKA signaling pathway; phosphorylation; cancer;
KW   kinase substrate; immunosuppressive disorder; proliferative disease;
KW   HIV infection; AIDS; immunodeficiency; autoimmune disease;
KW   systemic lupus erythematosus; Src-family.
XX
OS   Homo sapiens.
XX
PN   WO9962315-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-GB01680.
XX
PR   27-MAY-1998;   98NO-0002419.
PR   30-DEC-1998;   98US-0114240.
XX
PA   (LAUR-) LAURAS AS.
PA   (JONE/) JONES E L.
XX
PI   Hansson V,  Levy FO,  Mustelin T,  Skalhegg BS,  Sundvold V,  Tasken K;
PI   Vang T,  Altman A,  Munshi A;
XX
DR   WPI; 2000-086801/07.
DR   N-PSDB; AAZ46491.
XX
PT   Altering the activity of protein kinase signaling pathways, used for
PT   treating immunosuppressive disorders, e.g. AIDS, proliferative
PT   disorders, e.g. cancers or autoimmune diseases  -
XX
PS   Claim 23; Page 95-96; 111pp; English.
XX
CC   The invention provides a novel method of altering the activity of the
```

CC protein kinase A (PKA) signaling pathway in a cell that comprises
CC altering the extent of phosphorylation of one or more PKA substrates, or
CC kinase substrates downstream in the PKA signaling pathway. Pharmaceutical
CC compositions containing a nucleic acid molecule that encodes a PKA
CC substrate, or fragment, precursor or functionally equivalent variant,
CC where the sequence is modified to alter its susceptibility to
CC phosphorylation by PKA can be used for treating a disorder exhibiting
CC abnormal PKA signaling activity, immunosuppressive disorders or
CC proliferative diseases. They can be used for treating e.g. HIV
CC infection, AIDS, common variable immunodeficiency or cancers. Conditions
CC in which upregulation of the PKA pathway is required, such as autoimmune
CC disease, e.g. systemic lupus erythematosus, may also be treated. The
CC present sequence represents a PKA substrate, wherein the substrate is in
CC the Src-family, preferably Lck, Fyn, Src, Yes, Fgr, Lyn, Hck Blk, Yrk,
CC c-tkl, Fyk, Src-1 or Src-2.
XX
SQ Sequence 509 AA;

Query Match 26.2%; Score 352; DB 21; Length 509;
Best Local Similarity 41.7%; Pred. No. 3.7e-27;
Matches 75; Conservative 26; Mismatches 69; Indels 10; Gaps 2;

```
Qy      25 VTMEAERSKAT------AVALGSFPAGGPAELSLRLGEPLTIVSEDGDWWTVLSEVSGRE 78
           || |     |:       :|| |:      :|    |||| |: : |:||   | ·:|:|
Db      49 VTYEGSNPPASPLQDNLVIALHSYEPSHDGDLGFEKGEPLRILEQSGEWWKAQSLTTGQE 108

Qy      79 YNIPSVHVAKVS----HGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSVR 134
             ||   ||| :      | :: |||: ||  || |||  |:||||||::  ||:|||||
Db     109 GFIPFNFVAKANSLEPEPWFFKNLSRKDAERQLLAPGNTHGSFLIRESESTAGSFSLSVR 168

Qy     135 LSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPCVLQR 194
             : ::||:|  ||||  |||||:||| |  || ||:  :| :|  |  ||  |:
Db     169 DFDQNQGEVVKHYKIRNLDNGGFYISPRITFPGLHELVRHYTNASDGLCTRLSRPCQTQK 228
```

RESULT 9
AAB37700
ID AAB37700 standard; protein; 508 AA.
XX
AC AAB37700;
XX
DT 02-MAR-2001 (first entry)
XX
DE Human lymphocyte kinase.
XX
KW Human; lymphocyte kinase; protein co-ordinate data; lck; crystal.
XX
OS Homo sapiens.
XX
PN WO200070030-A1.
XX
PD 23-NOV-2000.
XX
PF 19-MAY-2000; 2000WO-US13881.
XX
PR 19-MAY-1999; 99US-0134965.
XX
PA (KINE-) KINETIX PHARM INC.
XX
PI Zhu X;
XX
DR WPI; 2000-687708/67.
XX
PT Crystal of a protein-ligand complex for identifying kinase inhibitors,
PT comprises a truncated lymphocyte kinase and a ligand, and diffracts
PT X-rays to determine atomic coordinates at a resolution greater than 5
PT angstroms -
XX
PS Claim 1; Page 434-5; 438pp; English.
XX
CC The present invention relates to a crystal of a protein-ligand complex
CC comprising a truncated lymphocyte kinase (lck) and a ligand. The crystal
CC diffracts X-rays so that the atomic coordinates of the protein-ligand
CC complex can be determined to a resolution of greater than 5.0 Angstroms.
CC The truncated lck used in the present invention comprises the globular
CC core of the corresponding full-length lck. The present sequence is the
CC full-length human lck protein. The crystal of the present invention may
CC be used to identify kinase inhibitors in screening assays, in drug
CC screening and drug design processes, to design, select or test inhibitors
CC of kinase enzymes, where the inhibitors are used as therapeutics for the
CC treatment and modulation of diseases, disease symptoms or the effect of
CC other physiological events mediated by kinases, having one or more kinase
CC enzymes involved in their pathology.
XX
SQ Sequence 508 AA;

Query Match 25.6%; Score 344; DB 21; Length 508;
Best Local Similarity 41.1%; Pred. No. 2.4e-26;
Matches 74; Conservative 26; Mismatches 70; Indels 10; Gaps 2;

```
Qy      25 VTMEAERSKAT------AVALGSFPAGGPAELSLRLGEPLTIVSEDGDWWTVLSEVSGRE 78
           || |     |:       :|| |:      :|    || | |: : |:||   | :|:|
Db      48 VTYEGSNPPASPLQDNLVIALHSYEPSHDGDLGFEKGEQLRILEQSGEWWKAQSLTTGQE 107

Qy      79 YNIPSVHVAKVS----HGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSVR 134
             ||   ||| :      | :: |||: ||  || |||  |:||||||::  ||:|||||
Db     108 GFIPFNFVAKANSLEPEPWFFKNLSRKDAERQLLAPGNTHGSFLIRESESTAGSFSLSVR 167

Qy     135 LSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPCVLQR 194
             : ::||:|  ||||  |||||:||| |  || ||:  :| :|  |  ||  |:
Db     168 DFDQNQGEVVKHYKIRNLDNGGFYISPRITFPGLHELVRHYTNASDGLCTRLSRPCQTQK 227
```

RESULT 10
ABG05994
ID ABG05994 standard; Protein; 70 AA.
XX
AC ABG05994;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #5985.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS70181.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 36353; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags

```
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   70 AA;

  Query Match             25.4%;  Score 342;  DB 22;  Length 70;
  Best Local Similarity   98.5%;  Pred. No. 1.8e-27;
  Matches   64;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      62 SEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIRE 121
           |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       6 SKDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSREKAEELLLLPGNPGGAFLIRE 65

Qy     122 SQTRR 126
           |||||
Db      66 SQTRR 70


RESULT 11
ABB57957
ID   ABB57957 standard; Protein; 517 AA.
XX
AC   ABB57957;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster polypeptide SEQ ID NO 663.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   N-PSDB; ABL02060.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Disclosure; SEQ ID NO 663; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   517 AA;

  Query Match             25.0%;  Score 336;  DB 22;  Length 517;
  Best Local Similarity   38.0%;  Pred. No. 1.7e-25;
  Matches   95;  Conservative   38;  Mismatches   79;  Indels   38;  Gaps    9;

Qy      10 LPSPSLSSSVQGQGPVTM-------EAERSKATA---VALGSFPAGGPAELSLRLGEPLT 59
           :| || :    || |  :       |:| : | |   |||  : |    :|| | || |
Db      36 IPMPSHA----GQPPEQIRPVPQIPESETAGANAKIFVALYDYDARTDEDLSFRKGEHLE 91

Qy      60 IVSE-DGDWWTVLSEVSGREYNIPSVHVAKV----SHGWLYEGLSREKAEELLLLPGNPG 114
           |::: ||||   |: :  |  ||| :|||:    :  | :  : | :||: |||| |
Db      92 ILNDTQGDWWLARSKKTRSEGYIPSNYVAKLKSIEAEPWYFRKIKRIEAEKKLLLPENEH 151

Qy     115 GAFLIRESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYISPRLTFPSLQALVDH 174
           |||||:|::|   ||||||       | ::|||  || |  :|: | || :|| ||:|
Db     152 GAFLIRDSESRHNDYSLSVR-----DGDTVKHYRIRQLDEGGFFIARRTTFRTLQELVEH 206

Qy     175 YSELADDICCLLKEPCVLQRAGPLPGKDIPLPVT---VQRTPLNWKELDSSLLF-SEAAT 230
           ||: :| :|  | :|||           |  |||     ||   |:  :|| |  :  :
Db     207 YSKDSDGLCVNLCKPCV----------QIEKPVTEGLSHRTRDQWEIDRTSLKFVRKLGS 256

Qy     231 GEESLLSEGL 240
           |:   : |||
Db     257 GQFGDVWEGL 266


RESULT 12
AAU74614
ID   AAU74614 standard; Protein; 541 AA.
XX
AC   AAU74614;
XX
DT   09-APR-2002  (first entry)
XX
DE   Perinuclear theca 32 (PT32) associated tyrosine kinase, c-Yes.
XX
KW   Perinuclear theca 32; PT32; contraceptive; fertility;
KW   oocyte activation; vaccine; globozoospermy; spermiogenesis;
KW   spermatozoa; tyrosine kinase; c-Yes; immunocontraceptive;
KW   bovine; protein.
XX
OS   Bos sp.
XX
PN   WO200190185-A2.
XX
PD   29-NOV-2001.
XX
PF   25-MAY-2001; 2001WO-CA00738.
XX
PR   25-MAY-2000; 2000CA-2307128.
PR   25-MAY-2000; 2000US-206979P.
XX
PA   (TOOH ) UNIV QUEENS KINGSTON.
PA   (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI   Oko R,  Sutovsky P;
XX
DR   WPI; 2002-097644/13.
XX
PT   Isolated perinuclear theca 32 polypeptide that interacts with activated
PT   tyrosine kinase c-Yes, for enhancing fertility, treating/diagnosing
PT   diminished fertility and abnormal spermiogenesis and for providing
PT   contraception -
XX
```

```
PS   Example; Fig 10; 103pp; English.
XX
CC   The invention describes an isolated perinuclear theca 32 (PT32)
CC   polypeptide (I) which interacts with tyrosine kinase c-Yes. (I) is
CC   useful for: enhancing fertility in a mammal; treating globozoospermy, by
CC   expressing (I) in spermatozoa; inhibiting fertilisation, by introducing
CC   (I) or its antigenic fragment into a mammal to eilicit an immune
CC   response; enhancing the ability of round spermatids to activate oocytes;
CC   treating or diagnosing diminished fertility and abnormal spermiogenesis;
CC   in providing contraception; identifying contraceptive and
CC   fertility-enhancing agents. The polynucleotide is useful for producing
CC   (I) by recombinant techniques, as vaccine, as diagnostic reagents, and
CC   for chromosome identification. An antibody against (I) is useful in
CC   immunological assays, in immunocontraceptive methods, to identify cells
CC   expressing (I), and to purify (I) by affinity chromatography. A
CC   transgenic animal is useful as an animal model for studying human
CC   fertility and reproductive biology, and for screening compounds to
CC   identify modulators of oocyte activation. The use of (I) prevents the
CC   entry of components which are detrimental to embryonic development into
CC   the oocyte during oocyte activation with crude sperm extract and avoids
CC   the propagation of viruses such as HIV (human immunodeficiency virus) and
CC   SIV (simian immunodeficiency virus) carried in the sperm. This is the
CC   amino acid sequence of the src tyrosine kinase c-Yes which is naturally
CC   occurring in sperm perinuclear theca and important in development,
CC   described in the method of the invention.
XX
SQ   Sequence   541 AA;

  Query Match             23.8%;  Score 321;  DB 23;  Length 541;
  Best Local Similarity   31.8%;  Pred. No. 6.3e-24;
  Matches   92;  Conservative   43;  Mismatches  104;  Indels   50;  Gaps    9;

Qy       1 GSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFPAGGPAELSLRLGEPLTI 60
           |   |   ::|||  |:     | ||:         |||  : |    :|| : ||   |
Db      70 GGASSSFSAVPSPYPSTLT---GGVTV--------FVALYDYEARTTDDLSFKKGERFQI 118

Qy      61 VSE-DGDWWTVLSEVSGREYNIPSVHVAKV----SHGWLYEGLSREKAEELLLLPGNPGG 115
           ::  :||||   |  :|:   ||| :||      :  | :  : |: || ||| |||  |
Db     119 INNTEGDWWEARSIATGKTGYIPSNYVAPADSIQAEEWYFGKMGRKDAERLLLNPGNQRG 178

Qy     116 AFLIRESQTRRGSYSLSVRLSRPASWDRIR-----HYRIHCLDNGWLYISPRLTFPSLQA 170
            ||:|||:| :|:|||||:|      || :|     ||::  ||||  ||: |  | |||
Db     179 LFLVRESETTKGAYSLSIR-----DWDEVRGDNVKHYKLRKLDNGGYYITTRAQFESLQK 233

Qy     171 LVDHYSELADDICCLLKEPC---------VLQRAGPLPGKDIPLPVTVQR---------- 211
           || || | || :|  |   |         : : |  || : : | | : :
Db     234 LVKHYREHADGLCHKLTTVCPTVKPQTQGLAKDAWELPRESLRLEVKLGQGCFGEVWMGT 293

Qy     212 ----TPLNWKELDSSLLFSEAATGEESLLSEGLRESL-SFYISLNDEAV 255
               | :  | |    :  ||   |  :: :   : |   |  :::| :
Db     294 WNGTTKVAIKTLKPGTMMPEAFLQEAQIMKKLRHDKLVPLYAVVSEEPI 342


RESULT 13
ABG10302
ID   ABG10302 standard; Protein; 543 AA.
XX
AC   ABG10302;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #10293.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS74489.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 20; SEQ ID No 40661; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   543 AA;

  Query Match             23.7%;  Score 319.5;  DB 22;  Length 543;
  Best Local Similarity   29.3%;  Pred. No. 9e-24;
  Matches   93;  Conservative   45;  Mismatches  112;  Indels   67;  Gaps    9;

Qy       1 GSLPSRRKSLPSPSLSSSV----------QGQGPVTMEAERSKATAVALGSFPAG---- 45
           |: |:     || |   :           |   ||     | : :|   |:|||
Db      33 GAEPTTVSPCPSSSAKGTAVNFSSLSMTPFGGSSGVTPFGGASSSFSVVPSSYPAGLTGG 92

Qy      46 ------------GPAELSLRLGEPLTIVSE-DGDWWTVLSEVSGREYNIPSVHVAKV-- 89
                       :|| : ||   |::  :||||   |  :|:   ||| :||
Db      93 VTIFVALYDYEARTTEDLSFKKGERFQIINNTEGDWWEARSIATGKNGYIPSNYVAPADS 152

Qy      90 --SHGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSVRLSRPASWDRIR-- 145
             :  | :  : |: || ||| ||| | ||:|||:| :|:|||:|      || ||
Db     153 IQAEEWYFGKMGRKDAERLLLNPGNQRGIFLVRESETTKGAYSLSIR-----DWDEIRGD 207

Qy     146 ---HYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPC---------VLQ 193
              ||:|  ||||  ||: |  | :|| || ||:| || :|  |   |          : :
Db     208 NVKHYKIRKLDNGGYYITTRAQFDTLQKLVKHYTEHADGLCHKLTTVCPTVKPQTQGLAK 267

Qy     194 RAGPLPGKDIPLPVTVQR--------------TPLNWKELDSSLLFSEAATGEESLLSEG 239
           |  :| : : | | : :                | :  | |     :  ||    |  :: :
Db     268 DAWEIPRESLRLEVKLGQGCFGEVWMGTWNGTTKVAIKTLKPGTMMPEAFLQEAQIMKKL 327

Qy     240 LRESL-SFYISLNDEAV 255
             : |   |  :::| :
Db     328 RHDKLVPLYAVVSEEPI 344


RESULT 14
AAB84663
```

```
ID   AAB84663 standard; Protein; 543 AA.
XX
AC   AAB84663;
XX
DT   05-SEP-2001  (first entry)
XX
DE   Amino acid sequence of human tyrosine kinase protein Yes.
XX
KW   Vascular permeability; tyrosine kinase protein; Src; Yes; stroke;
KW   myocardial infarction; restenosis; trauma; blood vessel; atherosclerosis;
KW   diabetic retinopathy; inflammatory disease; infection; arthritis;
KW   adult respiratory distress syndrome; ARDS; rheumatoid arthritis;
KW   diabetic retinopathy; psoriasis; neovascular glaucoma;
KW   capillary proliferation; osteoporosis; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200145751-A1.
XX
PD   28-JUN-2001.
XX
PF   22-DEC-2000; 2000WO-US35396.
XX
PR   22-DEC-1999;   99US-0470881.
PR   29-MAR-2000; 2000US-0538248.
XX
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Cheresh DA,  Eliceiri B,  Paul R;
XX
DR   WPI; 2001-417982/44.
DR   N-PSDB; AAH28359.
XX
PT   Modulating vascular permeability in tissues, including inflamed tissue,
PT   tissues associated with stroke, myocardial infarction, by contacting
PT   the tissue with tyrosine kinase protein Src, Yes or their modified
PT   forms -
XX
PS   Disclosure; Fig 11; 133pp; English.
XX
CC   The specification describes a method for modulating vascular
CC   permeability in a tissue suffering from a disease condition. The method
CC   comprises contacting the tissue with a pharmaceutical composition
CC   comprising tyrosine kinase protein Src, Yes or their mixtures or
CC   nucleic acid expressing them. The method is useful for modulating
CC   vascular permeability in tissues, including inflamed tissue, tissues
CC   associated with stroke, myocardial infarction or other blockage of
CC   normal flow, tissues undergoing restenosis, psoriatic, retinal tissue
CC   and similar tissues. Pathologies which may be treated include include
CC   trauma to blood vessels, and other systemic pathological events such as
CC   atherosclerosis, diabetic retinopathy, inflammatory disease due to
CC   infection by microbial agents and arthritis. Other diseases which can
CC   be treated include adult respiratory distress syndrome (ARDS), rheumatoid
CC   arthritis, diabetic retinopathy, psoriasis, neovascular glaucoma,
CC   capillary proliferation in atherosclerotic plaques and osteoporosis and
CC   cancer associated disorders such as solid tumours, solid tumour
CC   metastases, angiofibromas and hemangiomas. The present sequence
CC   represents human Yes, and is used in the method of the invention.
XX
SQ   Sequence   543 AA;

  Query Match           23.7%;  Score 319.5;  DB 22;  Length 543;
  Best Local Similarity   29.3%;  Pred. No. 9e-24;
  Matches   93;  Conservative   45;  Mismatches  112;  Indels   67;  Gaps    9;

Qy       1 GSLPSRRKSLPSPSLSSSV----------QGQGPVTMEAERSKATAVALGSFPAG---- 45
           |: |:     || |  :             |   ||     | : :|   |:|||
Db      33 GAEPTTVSPCPSSSAKGTAVNFSSLSMTPFGGSSGVTPFGGASSSFSVVPSSYPAGLTGG 92

Qy      46 ------------GPAELSLRLGEPLTIVSE-DGDWWTVLSEVSGREYNIPSVHVAKV-- 89
                        :|| : ||   |::  :|||   |  :|:   ||| :||
Db      93 VTIFVALYDYEARTTEDLSFKKGERFQIINNTEGDWWEARSIATGKNGYIPSNYVAPADS 152

Qy      90 --SHGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSVRLSRPASWDRIR-- 145
             :  | :  : |: || ||| |||  | ||:|||:| :|:||||:|      || ||
Db     153 IQAEEWYFGKMGRKDAERLLLNPGNQRGIFLVRESETTKGAYSLSIR-----DWDEIRGD 207

Qy     146 ---HYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPC---------VLQ 193
               ||:|  ||||  ||: |  | :|| || ||:| || :|  |   |         : :
Db     208 NVKHYKIRKLDNGGYYITTRAQFDTLQKLVKHYTEHADGLCHKLTTVCPTVKPQTQGLAK 267

Qy     194 RAGPLPGKDIPLPVTVQR--------------TPLNWKELDSSLLFSEAATGEESLLSEG 239
           |  :| : : | | : :              | :  | |    :  ||   |  :: :
Db     268 DAWEIPRESLRLEVKLGQGCFGEVWMGTWNGTTKVAIKTLKPGTMMPEAFLQEAQIMKKL 327

Qy     240 LRESL-SFYISLNDEAV 255
             : |   |  :::| :
Db     328 RHDKLVPLYAVVSEEPI 344


RESULT 15
AAY24421
ID   AAY24421 standard; Protein; 543 AA.
XX
AC   AAY24421;
XX
DT   23-SEP-1999  (first entry)
XX
DE   Human yes1 protein.
XX
KW   Human; yes1; diagnosis; neuropsychiatric disorder; BAD; schizophrenia;
KW   bipolar affective disorder; attention deficit disorder;
KW   schizoaffective disorder; unipolar affective disorder;
KW   Huntington's disease; Parkinson's disease; manic-depression.
XX
OS   Homo sapiens.
XX
PN   WO9935290-A1.
XX
PD   15-JUL-1999.
XX
PF   07-JAN-1999;   99WO-US00297.
XX
PR   08-JAN-1998;   98US-0003944.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Chen H,  Freimer NB;
XX
DR   WPI; 1999-444203/37.
DR   N-PSDB; AAX90200.
XX
PT   Detection of a genetic mutation in the yes1 gene, useful for
PT   diagnosis of a yes1 mediated neuropsychiatric disorder in a human
XX
PS   Disclosure; Fig 1; 110pp; English.
XX
CC   The present invention describes a method for detecting a genetic
CC   mutation in the yes1 gene for the diagnosis of a yes1 mediated
CC   neuropsychiatric disorder in a human. The method comprises detecting the
CC   presence or absence of a genetic mutation in the yes1 gene of the
CC   subject, where the genetic mutation is a substitution, insertion or a
CC   deletion and results in the production of a yes1 protein having an amino
CC   acid sequence other than the wild-type yes1 amino acid sequence and the
CC   presence of the genetic mutation identifies a subject that has or is at
CC   risk for developing a yes1 mediated neuropsychiatric disorder. Compounds
CC   that bind to the yes1 protein, alter the amount of the protein, or alter
CC   the activity of the yes1 gene product, are useful for treating a yes1
CC   mediated neuropsychiatric disorder. The disorders include Huntington's
CC   disease, Parkinson's disease, and especially bipolar-affected disorder
CC   (BAD) also known as bipolar mood disorder (BP) or manic-depressive
CC   illness. The method distinguishes neuropsychiatric disorders from
CC   neurological disorders, which enables more accurate evaluation and
CC   prescription of medical treatment. The present sequence represents the
CC   human yes1 protein sequence.
XX
```

```
SQ   Sequence   543 AA;

  Query Match            23.6%;  Score 317.5;  DB 20;  Length 543;
  Best Local Similarity  29.3%;  Pred. No. 1.4e-23;
  Matches   93;  Conservative   45;  Mismatches  112;  Indels   67;  Gaps    9;

Qy       1 GSLPSRRKSLPSPSLSSSV-----------QGQGPVTMEAERSKATAVALGSFPAG---- 45
           |: |:     || |   :              |   ||      | : :|   |:|||
Db      33 GAEPTTVSPCPSSSAKGTAVNFSSLSMTPFGGSSGVTPFGGASSSFSVVPSSYPAGLTGG 92

Qy      46 ------------GPAELSLRLGEPLTIVSE-DGDWWTVLSEVSGREYNIPSVHVAKV-- 89
                        :|| : ||   |::  :||||   |  :|:   ||| :||
Db      93 VTIFVALYDYEARTTEDLSFKKGERFQIINNTEGDWWEARSIATGKNGYIPSNYVAPADS 152

Qy      90 --SHGWLYEGLSREKAEELLLLPGNPGGAFLIRESQTRRGSYSLSVRLSRPASWDRIR-- 145
             :  | :  : |: || ||| |||   | ||:|||:| :|:||||:|      || ||
Db     153 IQAEEWYFGKMGRKDAERLLLNPGNQRGIFLVRESETTKGAYSLSIR-----DWDEIRGD 207

Qy     146 ---HYRIHCLDNGWLYISPRLTFPSLQALVDHYSELADDICCLLKEPC---------VLQ 193
               ||:|  ||||  ||: |  | :|| || ||:| || :|  |   |         : :
Db     208 NVKHYKIRKLDNGGYYITTRAQFDTLQKLVKHYTEHADGLCHKLTTVCPTVKPQTQGLAK 267

Qy     194 RAGPLPGKDIPLPVTVQR--------------TPLNWKELDSSLLFSEAATGEESLLSEG 239
             |  :| : : | | : :              | :  | |    :  ||   |  :: :
Db     268 DAWEIPRESLRLEVKLGQGCFGEVWMGTWNGTTKVAIKTLCPGTMMPEAFLQEAQIMKKL 327

Qy     240 LRESL-SFYISLNDEAV 255
             : |   |  :::| :
Db     328 RHDKLVPLYAVVSEEPI 344


Search completed: March 24, 2003, 15:48:35
Job time : 57.1515 secs
```